**GIBS v3 | Sampling Plan: India**

|  |  |
| --- | --- |
| **GENERAL INFORMATION** |  |
| Country | India |
| Mode – substantive and contact (CAPI or PAPI) | CAPI |
| Targeted sample size | 1,500 internet user completes for the main sample and an additional 200 internet user completes among 15–24-year-olds. |
| Interview Languages | Hindi, Punjabi, Marathi, Bengali, Gujarati, Assamese, Odiya, Telugu, Kannada, Malayalam, Tamil |
| Coverage/Representativeness  \*This should coincide with calculations in the allocation table | The coverage rate is 93% of the target population (adults, non-institutionalized population). |
| Exclusions  \*Please specify source of data for calculation, if necessary | Excluded areas total approximately 7% of the population and include the north-eastern states, union territories, and islands: Andaman & Nicobar Islands, Arunachal Pradesh, Dadra and Nagar Haveli, Daman and Diu, Lakshadweep, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura. The aforementioned were excluded because of their extreme geographical location and the sparse population. Besides Jammu and Kashmir is also excluded due to the political and safety situation. |
| Data source for sample design  \*Clearly indicate which are used for PSU allocation, PSU frame, and SSU frame if differ | India Census 2011  Census 2011 is used for sample frame construction, allocation and sample selection. Sample is allocated based on population and selections are based on households. |
| Quota (if applicable) | There will be no demographic quotas |
|  |  |
| **STRATIFICATION** |  |
| General description | The design is stratified cluster sampling with 33 viable strata and 150 selected SSUs or ultimate clusters. |
| Number of strata (also to be included in an Excel allocation table) | Included units are stratified by States and urban/rural definition of settlements as defined by and marked in 2011 Census Lowest administrative units. A total of 24 states/territories are covered as part of this study. To avoid having too few completes allocated to a stratum, some states were not further split into urban/rural. Of the 24 states/territories covered, only 11 states have a large enough urban and rural population to have both urban and rural allocation, resulting in a total of 33 viable strata. In the remaining states/territories, units are implicitly stratified by urban/rural, by sorting the units. |
| Definition of strata | State by urbanity (urban vs rural) for those with a large enough urban population. And entire states for the remaining. |
| Sample allocation to strata (also to be included in an Excel allocation table | Proportional to 0+ population. |
|  |  |
| **SELECTION WITH CERTAINTY** |  |
| Number of PSUs selected with certainty due to size (if selected more than once, please note) | 0 |
| Number of PSUs selected with certainty due to importance | 0 |
| Number of PSUs selected with certainty due to being the only unit in the stratum | 0. |
| Selection with certainty taken into account in strata and sampling unit counts above? | NA |
|  |  |
| **OVERSAMPLES (IF APPLICABLE)** |  |
| Structure of oversamples | None |
|  |  |
| **STAGES OF SELECTION** |  |
| First stage of selection = PSUs (Primary Sampling Units) |  |
| Definition of PSU | Urban and rural parts of districts |
| Number of units in universe | 975 |
| Number of units in sampling frame | 839 |
| Number of units in sample | 120 |
| Selection method (e.g., PPS, SRS, purposive)  \*If PPS, include measure of size. If systematic, include sort order | PPS with number of Households as measure of size from 2011 census. |
| Number of interviews per unit |  |
| Proposed substitution procedure; weighting procedure for substitutes | Substitutions will be PSUs within the same strata and that are “most similar” to originally selected PSUs in terms of number of HHs. |
| Second stage of selection = SSUs (Secondary Sampling Units) |  |
| Definition of SSU | Wards in urban PSUs and Villages in Rural PSUs |
| Number of units in sampling frame (after exclusions) | 595720 |
| Number of units in sample | 150 |
| Selection method (e.g., PPS, SRS, purposive)  \*If PPS, include measure of size. If systematic, include sort order | PPS with number of HH as measure of size |
| Number of interviews per unit | 10 internet user completes for the main sample and an additional 1 to 2 internet user completes among 15-24 from the same cluster. |
| Proposed substitution procedure; weighting procedure for substitutes | Before sampling starts, the list of SSUs within selected PSU is generated by sorting the sampling frame in an increasing order within each stratum and PSU based on number of households at the village or ward level. Once SSU sampling is completed using PPS method, the selected SSUs is marked on the ordered list. When substitution is triggered due to unpredictable reasons, the alternative SSU within the selected PSU, is the unmarked one that is immediately preceding the selected within the same stratum and PSU on the sorted list. SSUs that are already sampled or used as a replacement are not chosen again for replacement.  The substitution procedure is designed to ensure that the alternative is from the same stratum and PSU as the one being replaced is “most similar” to originally selected SSUs in terms of number of HHs. |
| *Add additional stages of selection as needed* | *Please add a new section if you have the third or fourth stage of selection.* |
|  |  |
| **HOUSEHOLD AND RESPONDENT SELECTION** |  |
| Definition of household | Definition of a Household: All interviews will take place at a person’s home, which can be anything from a one-room flat to a single house. To be eligible, a household has to have its own cooking facilities, which could be anything from a standing stove in the kitchen to a small fire in the courtyard. |
| Household selection | Once at the given starting point, the interviewer will have to place his or her **back** to the (main) entrance of the structure and move to the right (Rule: *Always go to the right*). Regardless of whether there is one interviewer or more than one interviewer working in a starting point, the following instructions are valid. Counting **three** households (excluding the starting point), the interviewer will attempt a contact at the third household. The third HH is the 1st main HH, where the interviewer will make three attempts, on separate occasions, to secure an interview with an eligible household member. The third household could be a structure like an apartment building, in which case, the interviewer will enter the building continuing using the rules outlined later, to determine the household to be selected. Every effort should be made so that those attempts are distributed over more than one day to provide an opportunity for hard-to-reach individuals to be included in the sample. There may be times when it is not possible to return to an area on another day and attempts will have to be made on the same day. In this case, it is important to leave a gap of at least 2 hours between each attempt, unless it is by appointment. Interval of **three** is going to be implemented in rural and **five** in urban areas. Structures that are not occupied or where no one lives will not be counted as part of the interval.  Usually, each PSU is covered by one interviewer. On rare occasions, if the PSU is covered by two interviewers, each interviewer does one household at a time following the random route. The first interviewer approaches the first main household. Then after skipping 3 households, the second interviewer would approach the second main household, attempting to keep this order for the entire cluster. But, at times if someone finishes the interview first, they can continue with the random route procedure using the appropriate skip from the previously selected household until the required number of interviews are completed in that cluster.  Depending on the size of the locality and how many PSUs are assigned in that locality, an interviewer visits more than one PSU. For example, a small town with two PSUs will be walked through by one interviewer accompanied by a supervisor. If it is a town with many PSUs it is split between multiple interviewers accompanied by supervisors. |
| Household replacements | Any unsuccessful interviews at main households are compensated by attempts at replacement households. To identify the first replacement household, count three households past the last main household along the walking path. In each replacement household, up to 3 attempts are made using the same rules described for obtaining an interview with a main household. If attempts at 1st Replacement (R1) also result in failure, interviewer continue to use the right-hand rule and the interval of three to determine the next replacement household (R2). This process is repeated until the required number of completes is obtained in the selected PSU/Ultimate cluster. |
| Respondent selection method and eligibility | All adults 15 and older living permanently in the household are eligible.  Once at a given household, interviewers ask for the names, gender and ages of all individuals aged 15 and older who live permanently in the household, whether they are present or not at the time of the interviewer’s visit. Details of each household member is entered into the tablet, and the CAPI program randomly selects one individual to be interviewed. All eligible household members have an equal chance of being selected for the interview. If selected respondent is not available at the time of the visit, the interviewer needs to go again to the HH at a different time/day to give the opportunity to interview the selected respondent. If it is a third visit at that household, the interviewer cannot come back and needs to replace that household. In addition, the interviewer must code the reason for the replacement on the tracking sheet. Once an eligible respondent is selected, they complete the screening section and if qualified (internet user) proceed to complete the survey. If they fail the screener, basic demographic information such as age, gender, education and geography is captured. This process is repeated until the required number of internet user interviews are completed as part of the main sample. Once the main sample is completed, interviewers continue along the random route and only attempt to interview in households with at least one adult 15-24. The CAPI program is set up to change the screening criteria once the main sample is completed. |
| Number of contact attempts per sampled unit | In each sampled unit (PSU/SSU) interviewers complete 10 internet user interviews for the main sample and an additional 1-2 internet user interviews among 15-24. A maximum of 3 attempts per main and replacement households is made to complete the sample allocated to each cluster. |
| Timing of contact attempts (within a single day and across different days; minimum time between contact attempts if no appointment made; proportion of weekend days and evening hours for contact attempts; maximum number of contact attempts on any single day) | Attempts are made at different time of the day (at least 2 hours between each contact attempt unless an appointment is set with the respondent) and the third attempt happens on a different day of the week as the first two. One attempt at least is conducted after 5 pm or on a weekend. |
| Number of interviews allowed in single apartment building | There is not a limit of interviews to be completed in a building or multi-level dwelling. Interviewers continue using the interval of three until the required number of interviews are completed. |
| Incentives | No monetary or gifts are given to the respondents. Interviewers are trained to constantly use motivational techniques to increase participation. |
|  |  |